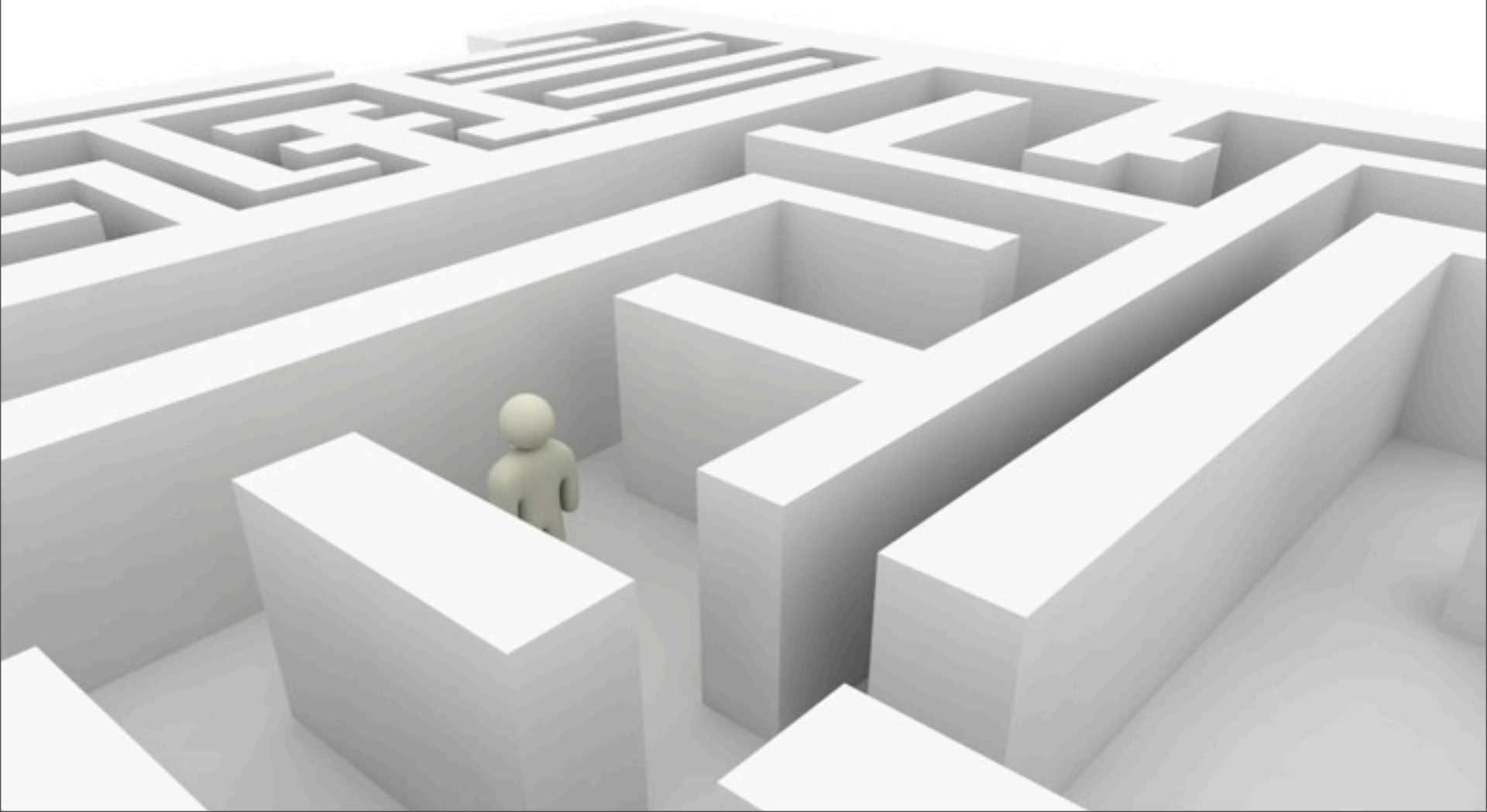
Software Quality
Prediction

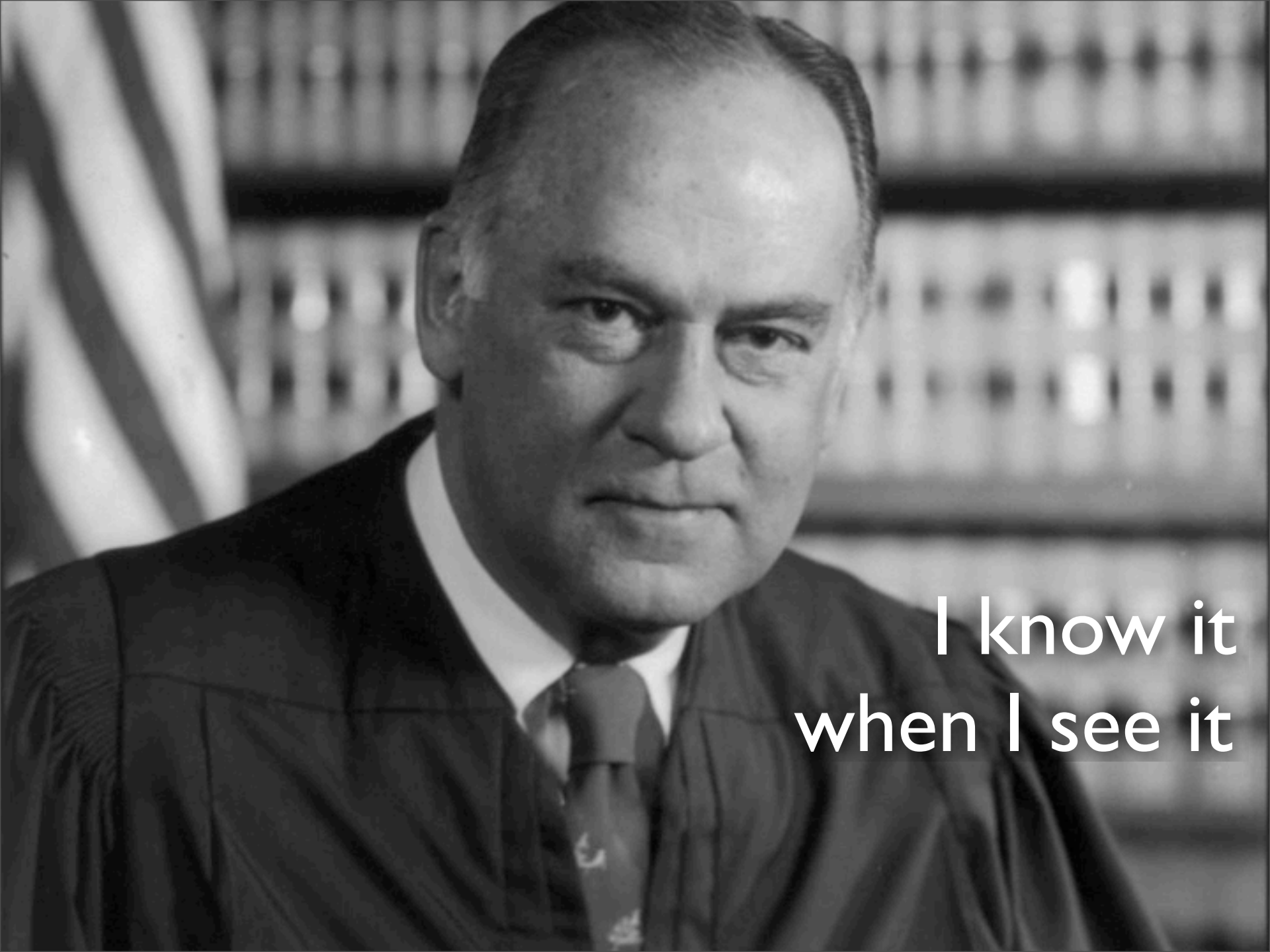
Stefan Wagner

The 15th CREST Open Workshop
25 October 2011
London, UK

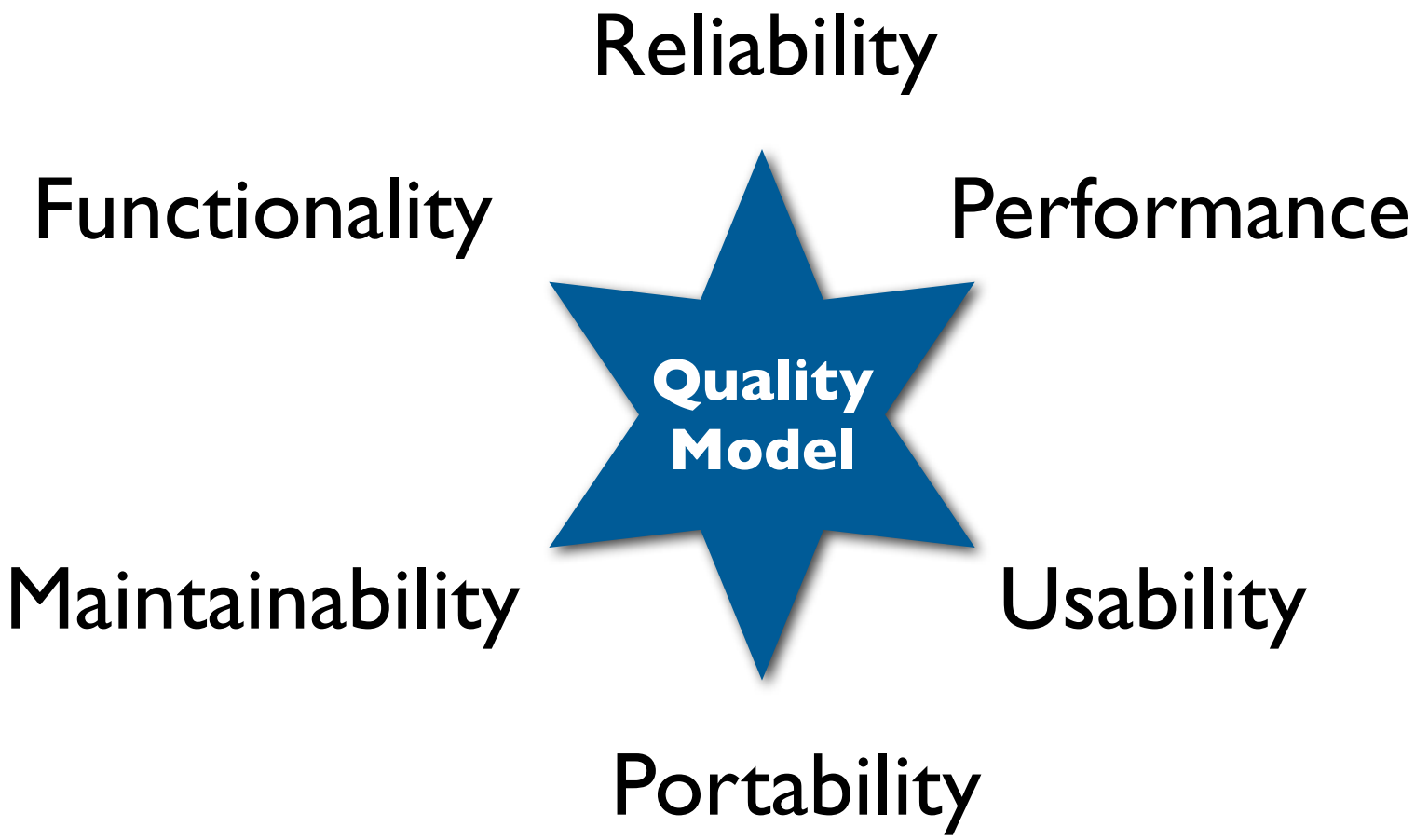
"Quality is a complex and multi-faceted concept...
it is also the source of great confusion."

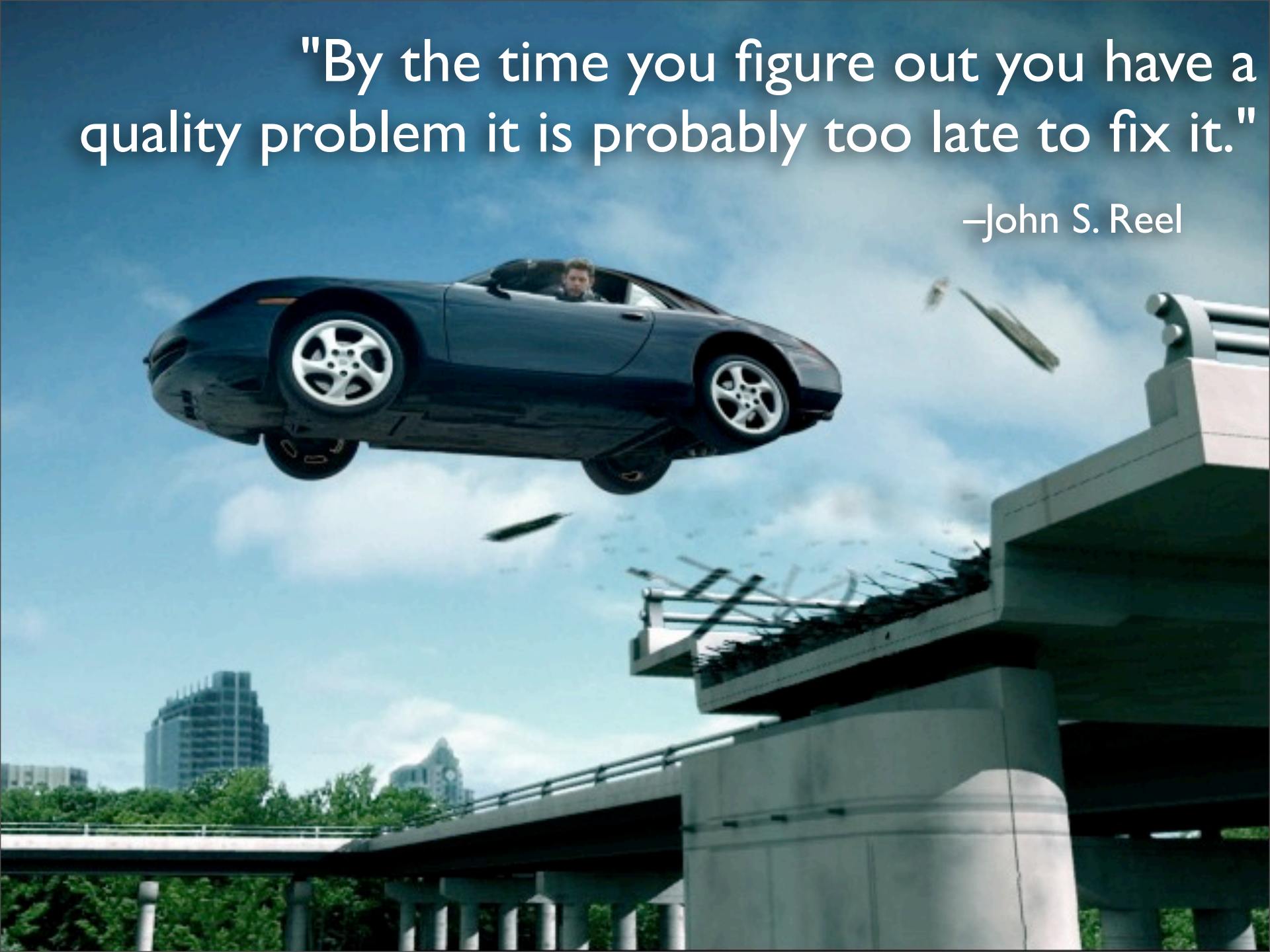
–David A. Garvin

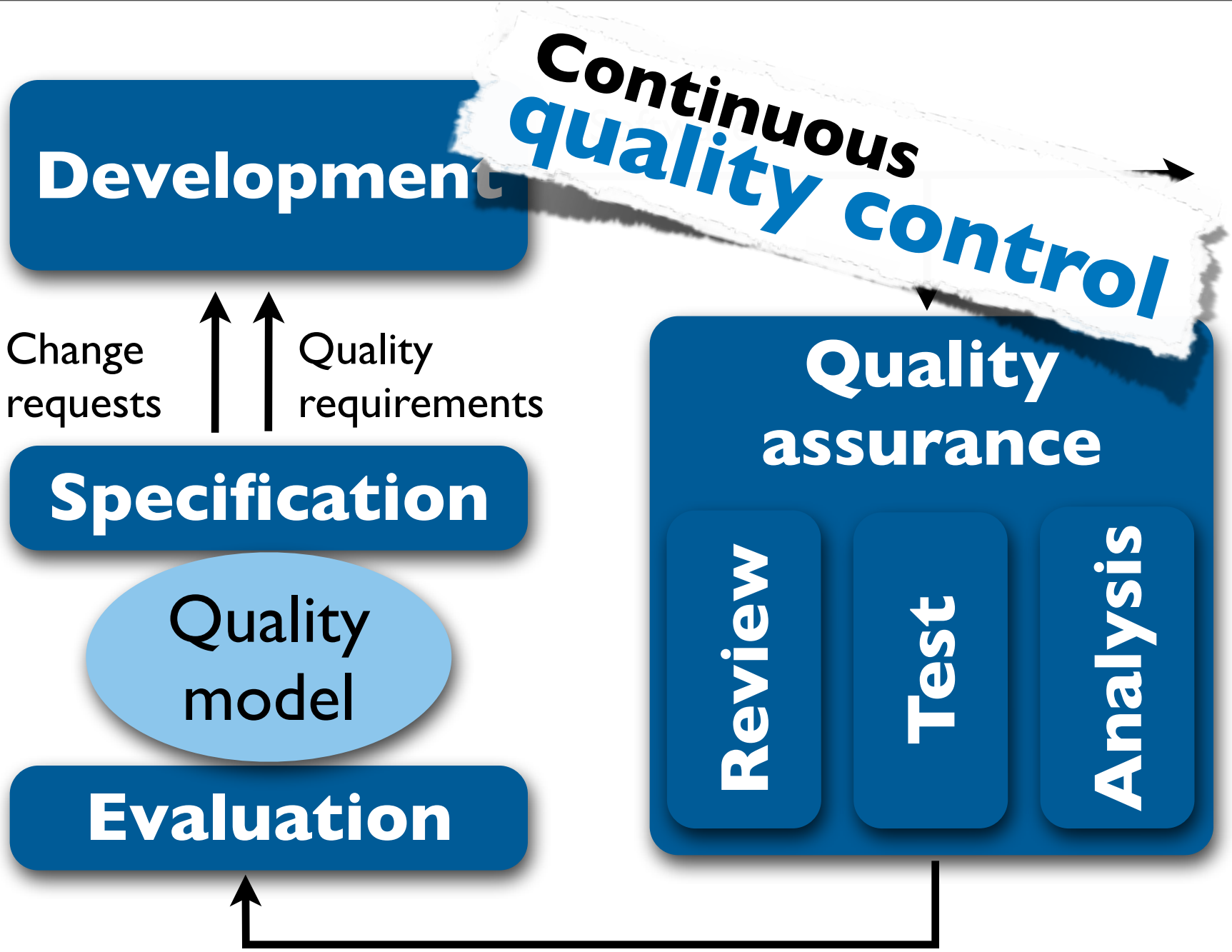



I know it
when I see it

ISO 9126

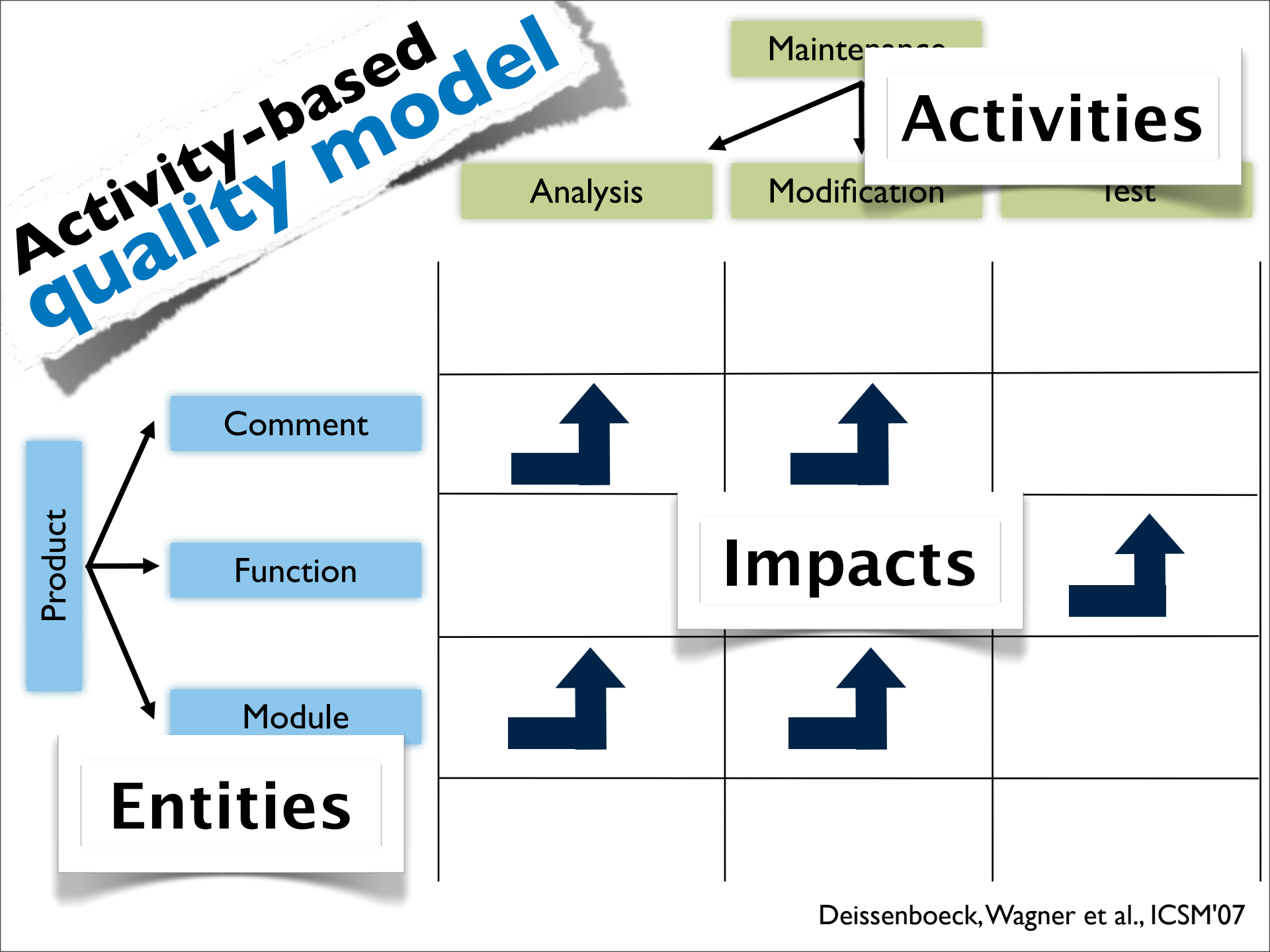

Reliability
Functionality
Performance
Quality Model
Maintainability
Usability
Portability


"By the time you figure out you have a quality problem it is probably too late to fix it."
–John S. Reel

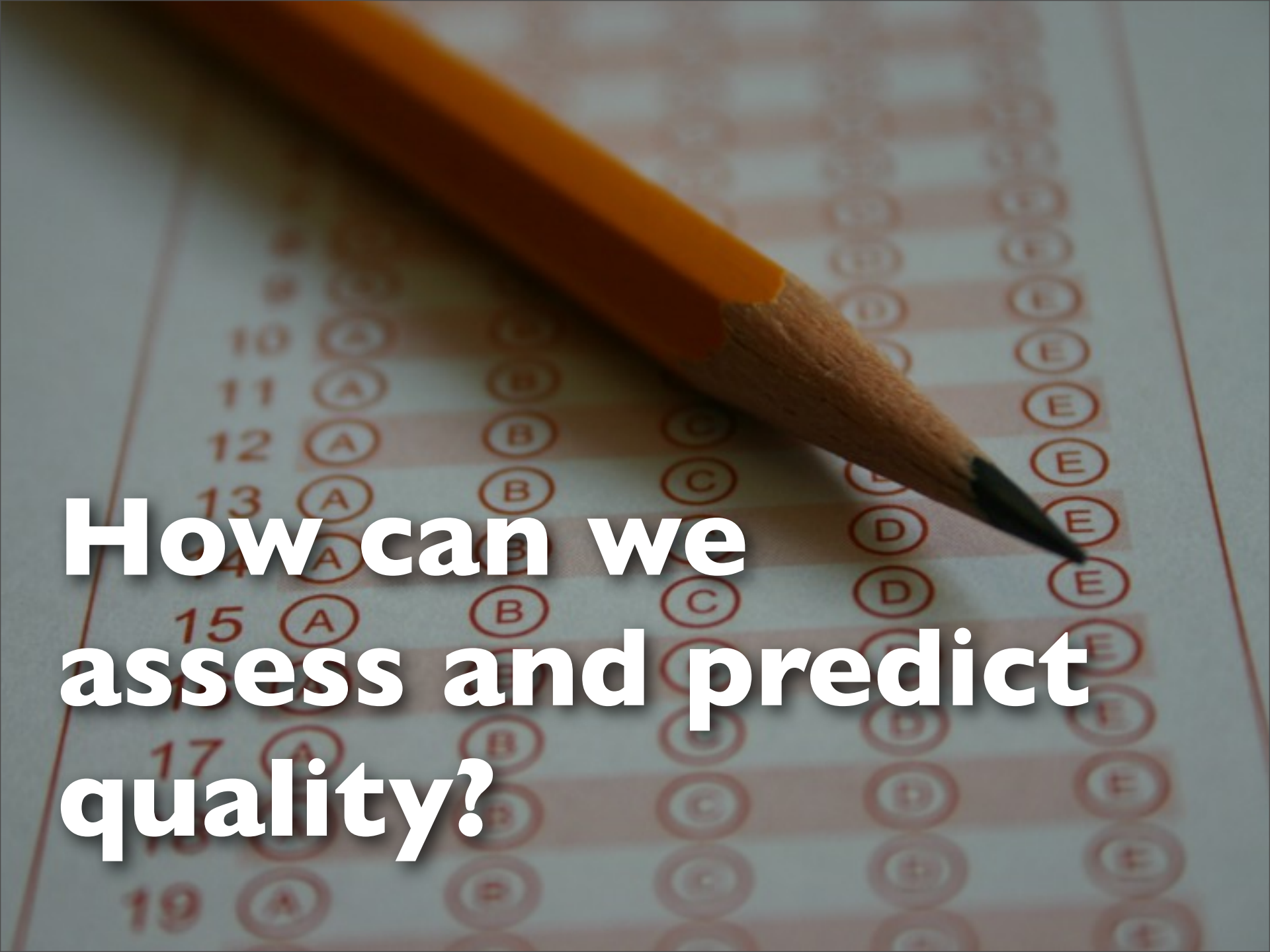


Deissenboeck, Wagner et al., IEEE Software, 2008

Activity-based quality model



Deissenboeck, Wagner et al., ICSM'07


How can we assess and predict quality?

Prediction research concentrates on bug prediction...

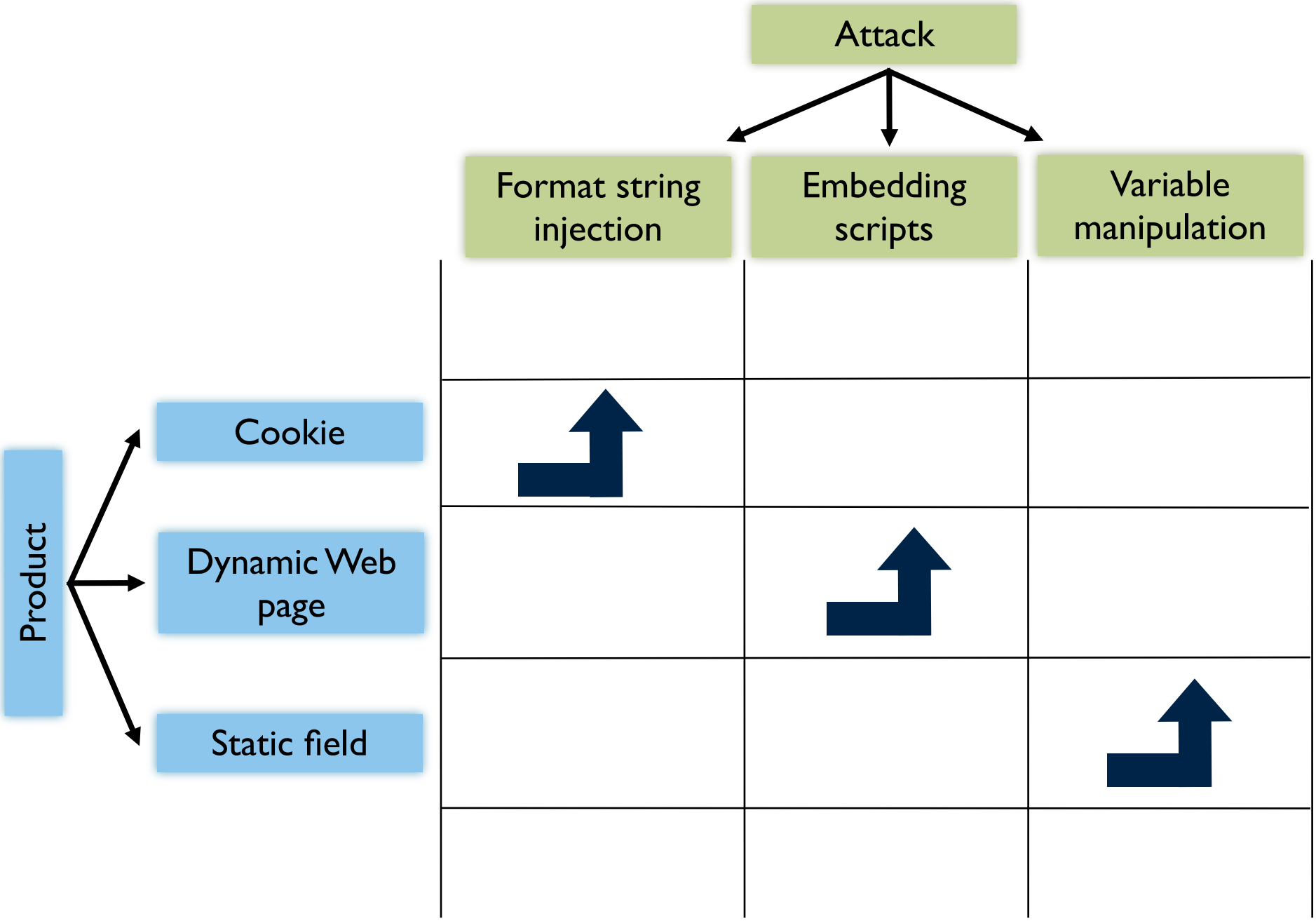
1. Scoring approach

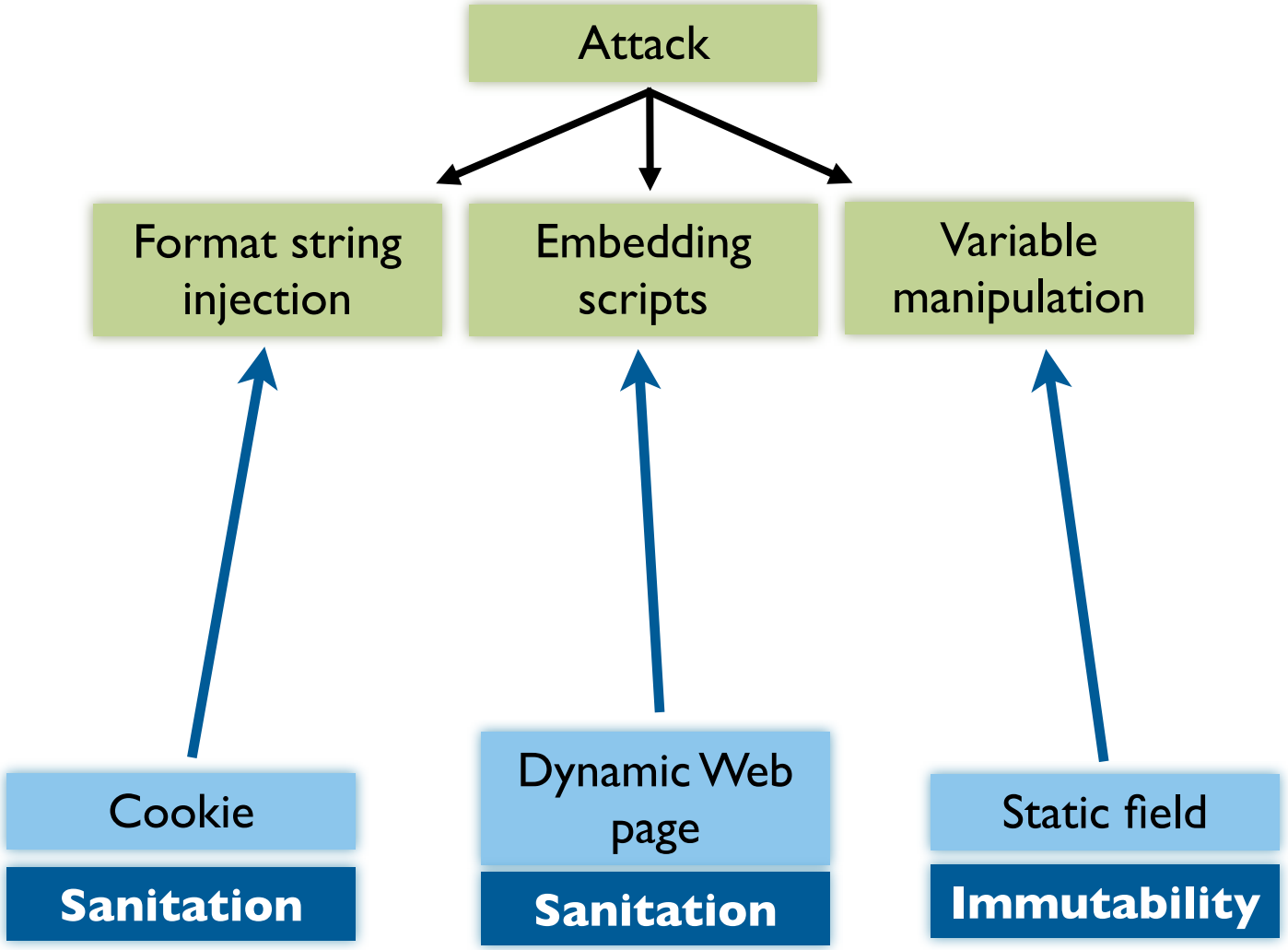
Quamoco

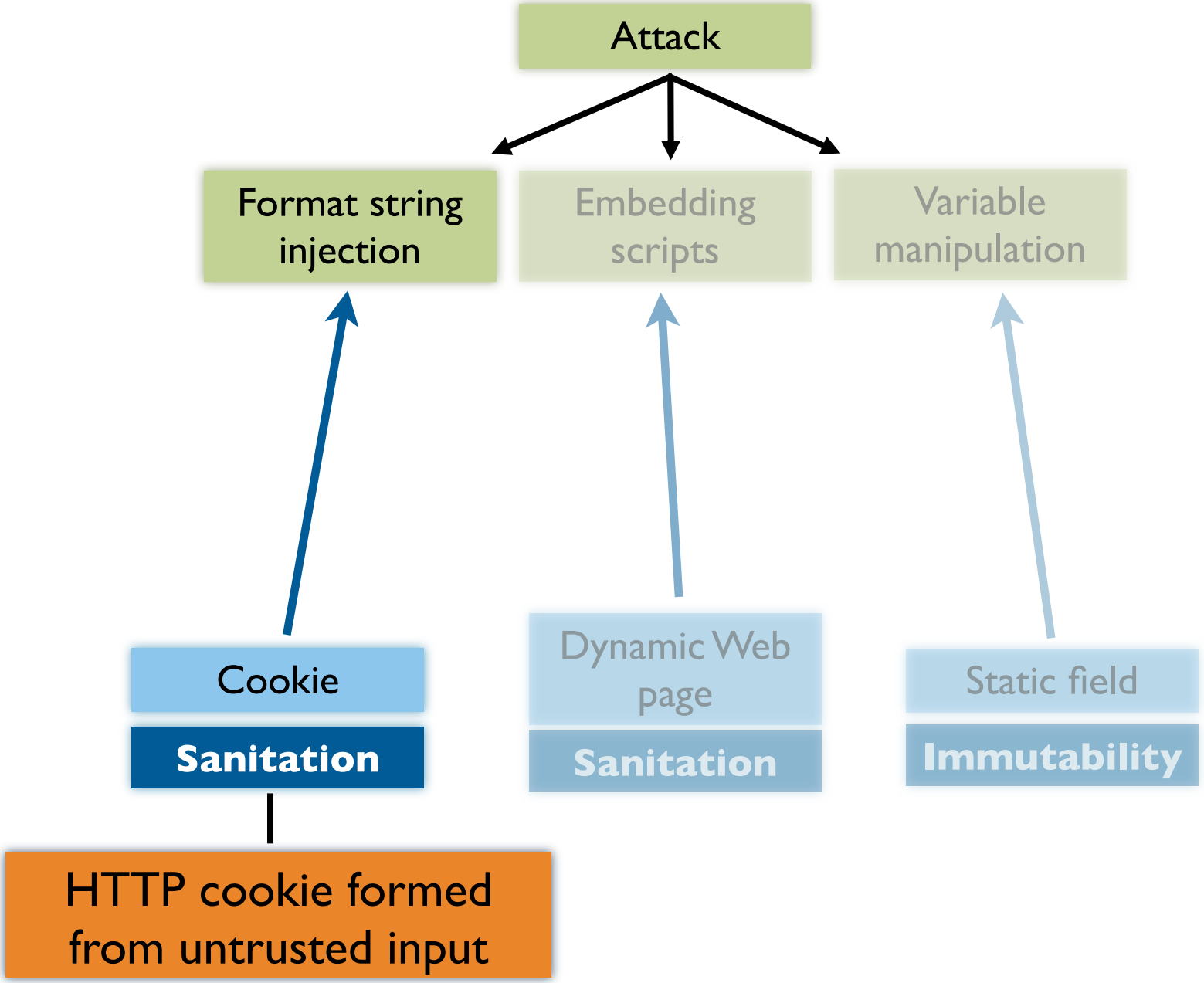
The Benchmark for Software Quality

Bundesministerium
für Bildung
und Forschung

Project partners


Attack
Format string injection
Embedding scripts
Variable manipulation
Product
Cookie
Dynamic Web page
Static field

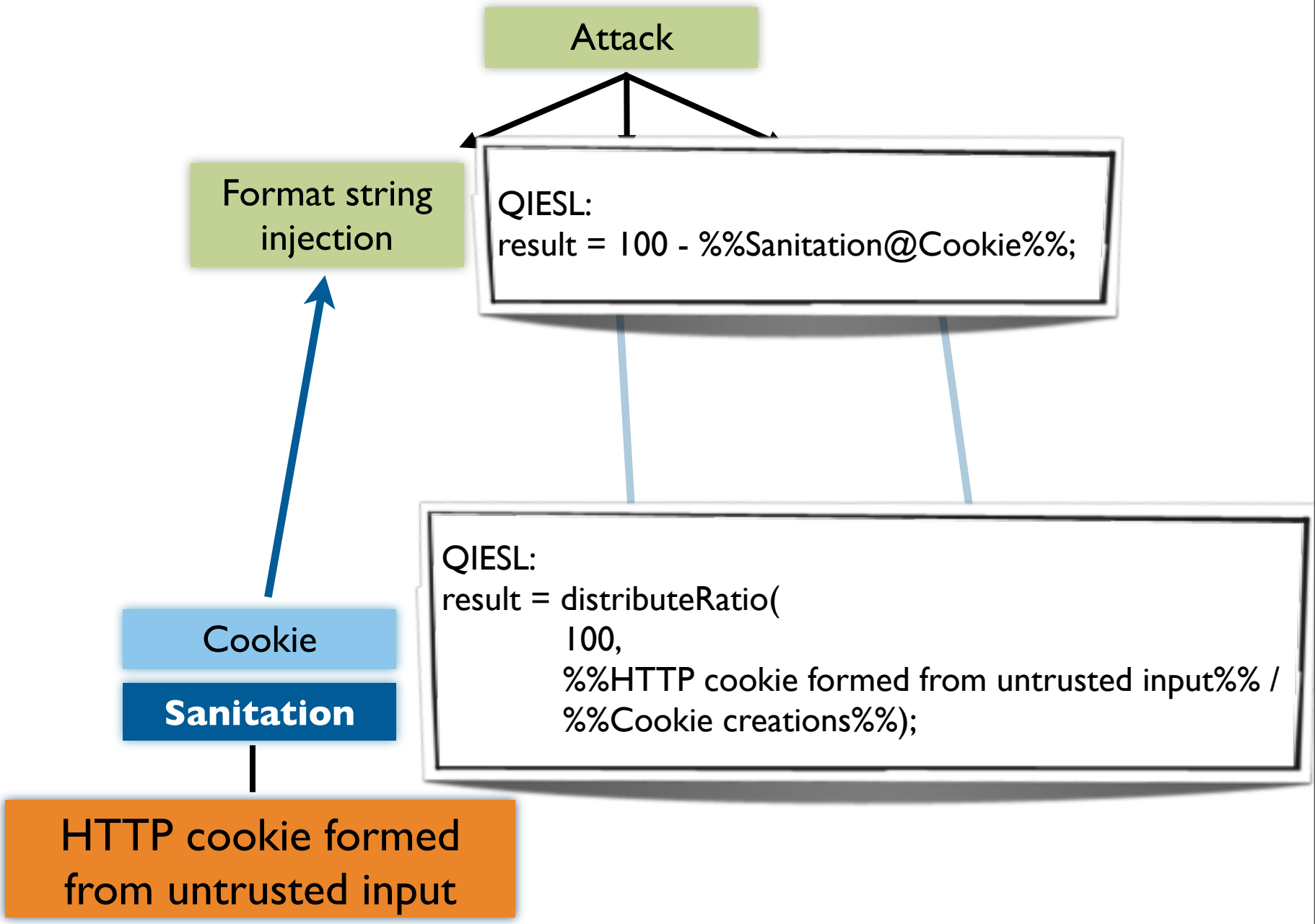

Attack
Format string injection
Embedding scripts
Variable manipulation
Cookie
Sanitation
Dynamic Web page
Sanitation
Static field
Immutability

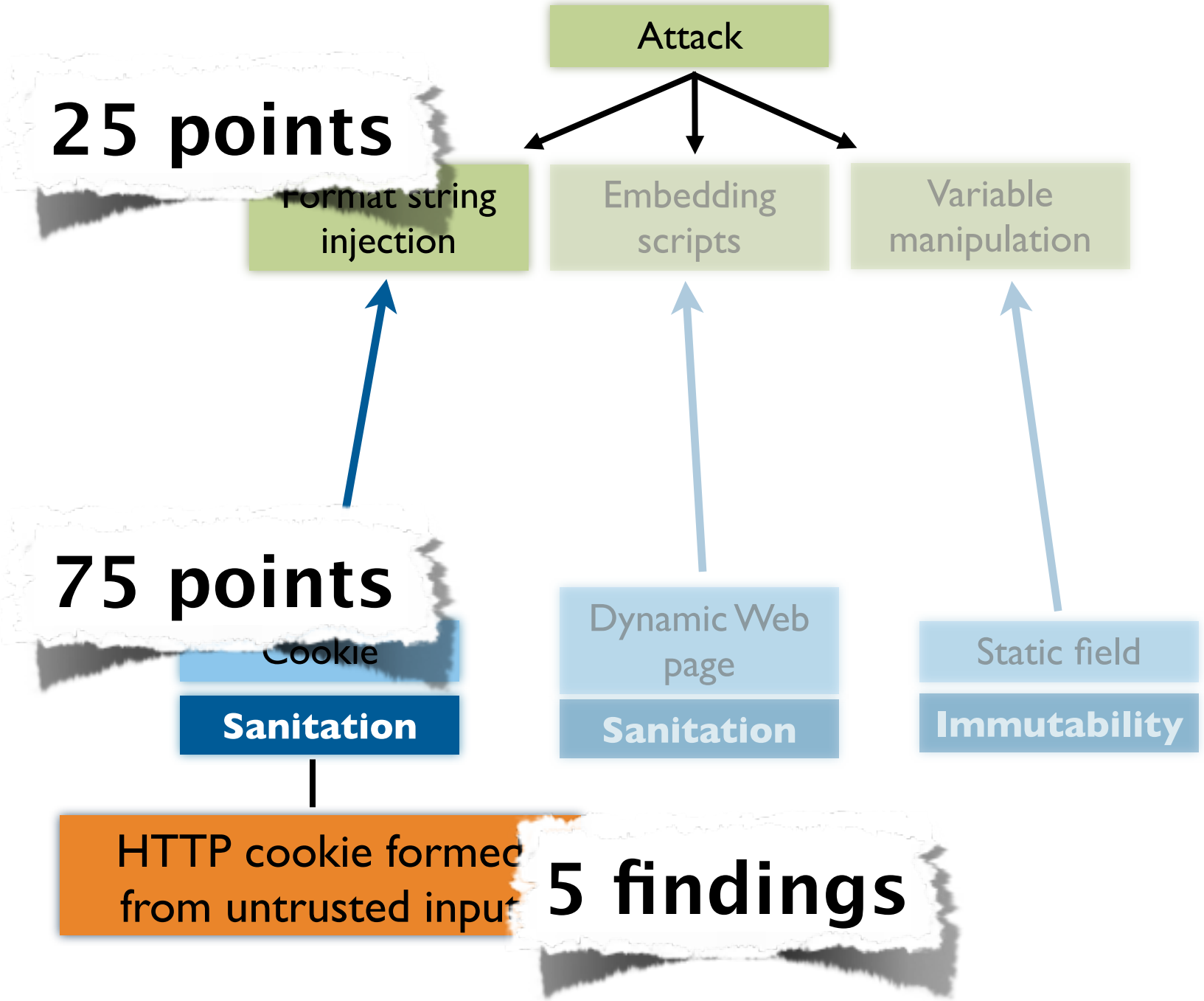

Attack
Format string injection
Embedding scripts
Variable manipulation
Cookie
Sanitation
Dynamic Web page
Sanitation
Static field
Immutability
HTTP cookie formed from untrusted input

Quality Impact Evaluation Specification Language (QIESL)

```
result = distributeRatio(
        100,
        %%Missing destructor%% /
        %%#Classes%%);
```

- Java-based syntax
- Access to factors and measures
- Helper functions
- Aggregation, evaluation, calibration


Attack
Format string injection
QIESL:
result = 100 - %%Sanitation@Cookie%%;
QIESL:
result = distributeRatio(
100,
%%HTTP cookie formed from untrusted input%% /
%%Cookie creations%%);
Cookie
Sanitation
HTTP cookie formed from untrusted input

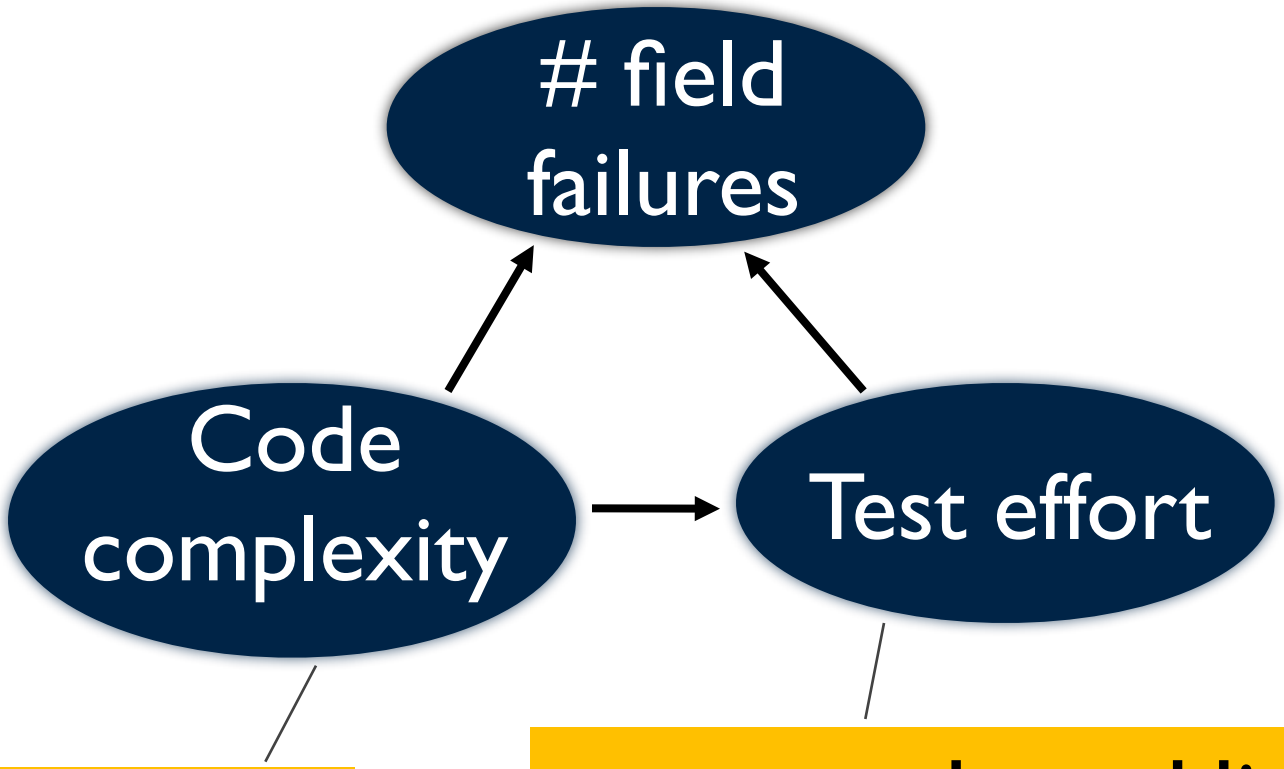

Attack
25 points
format string injection
Embedding scripts
Variable manipulation
75 points
Cookie
Sanitation
Dynamic Web page
Sanitation
Static field
Immutability
HTTP cookie formed from untrusted input
5 findings

Validation on OSS projects

	Ranking from model	Ranking from experts
Best	Checkstyle	Checkstyle
	Log4J	Log4J
	RSSOwl TV-Browser	RSSOwl
		TV-Browser
Worst	JabRef	JabRef

2. Bayesian net

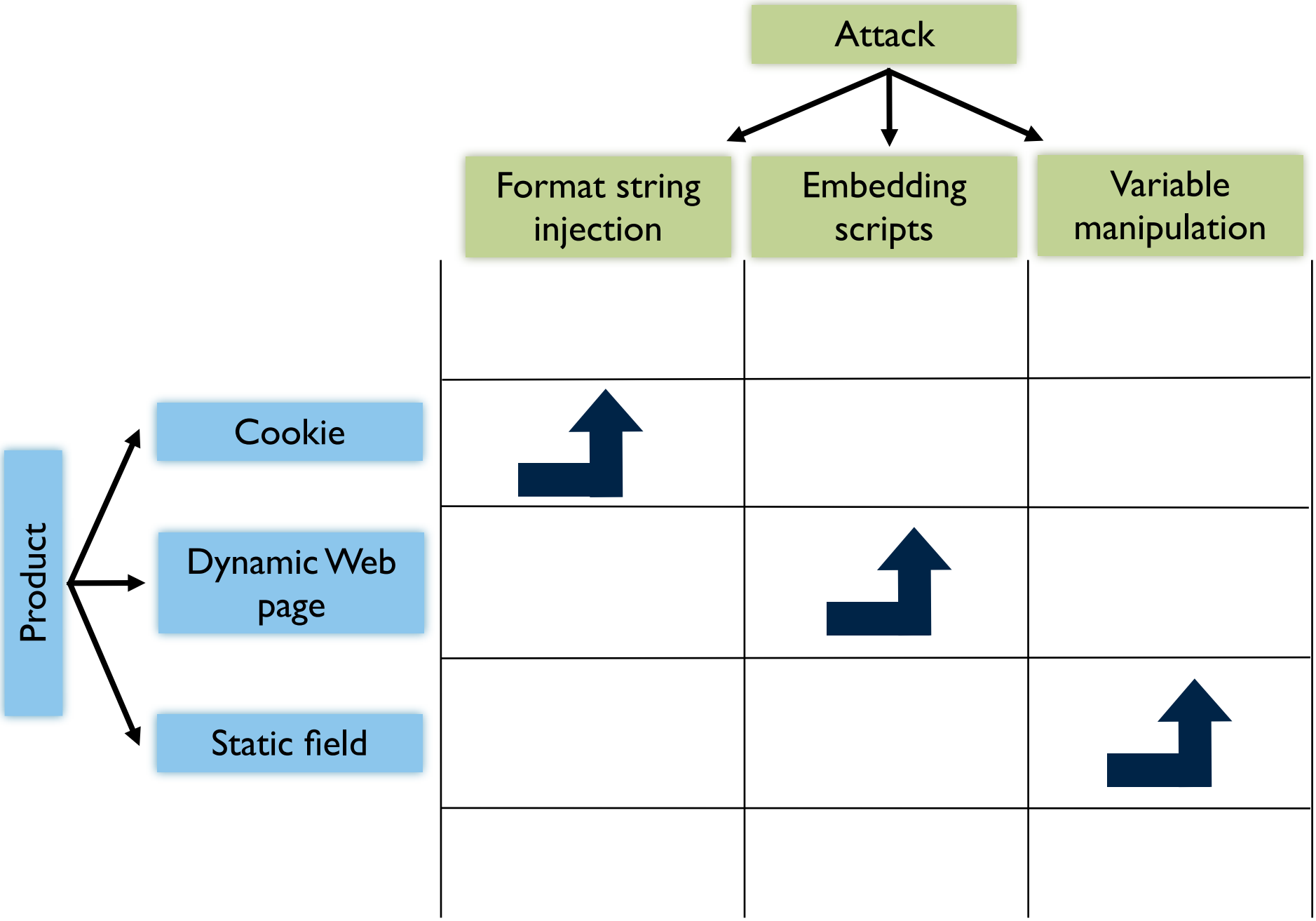
Bayesian net example

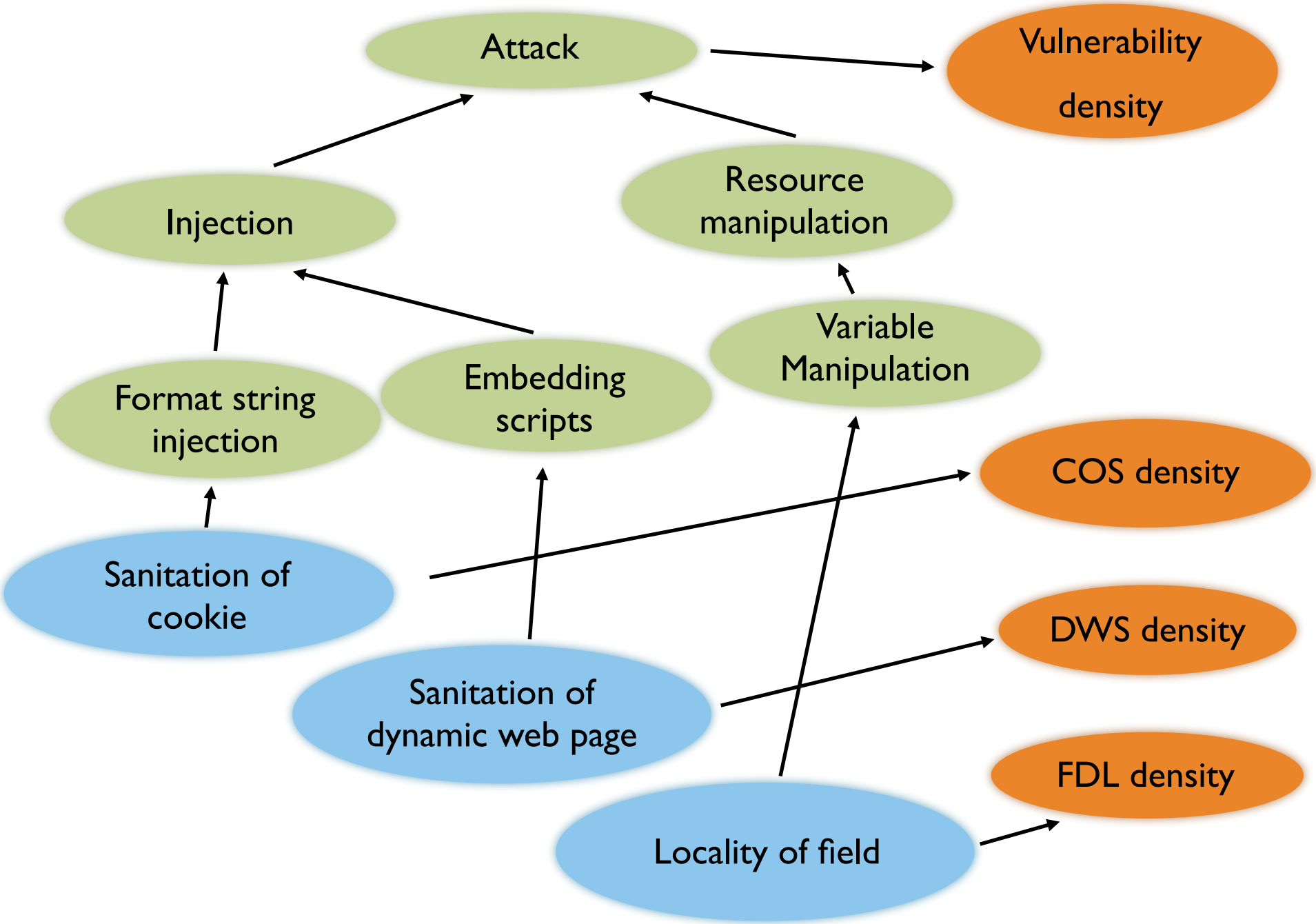


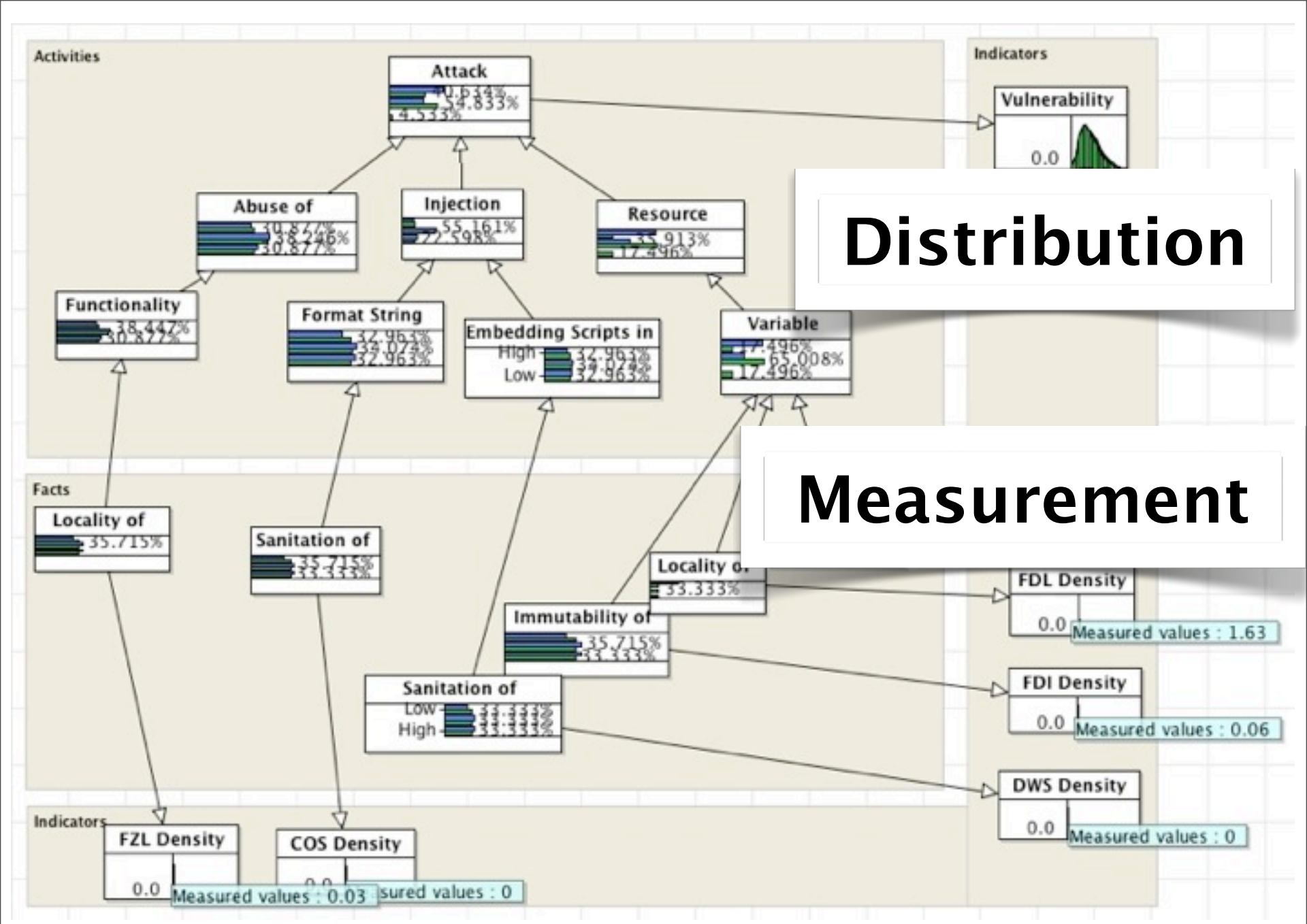
	Low = 0.6
	High = 0.4

Node Probability Table

	Low	High
Small	0.7	0.1
Med	0.2	0.2
Large	0.1	0.7


Attack
Format string injection
Embedding scripts
Variable manipulation
Product
Cookie
Dynamic Web page
Static field


Attack
Vulnerability density
Injection
Resource manipulation
Format string injection
Embedding scripts
Variable Manipulation
COS density
Sanitation of cookie
DWS density
Sanitation of dynamic web page
FDL density
Locality of field



AgenaRisk: http://www.agenarisk.com

Validation

Goals: gather experiences
test predictive validity

Maintainability

7 – 43 KLOC
3 – 6 years

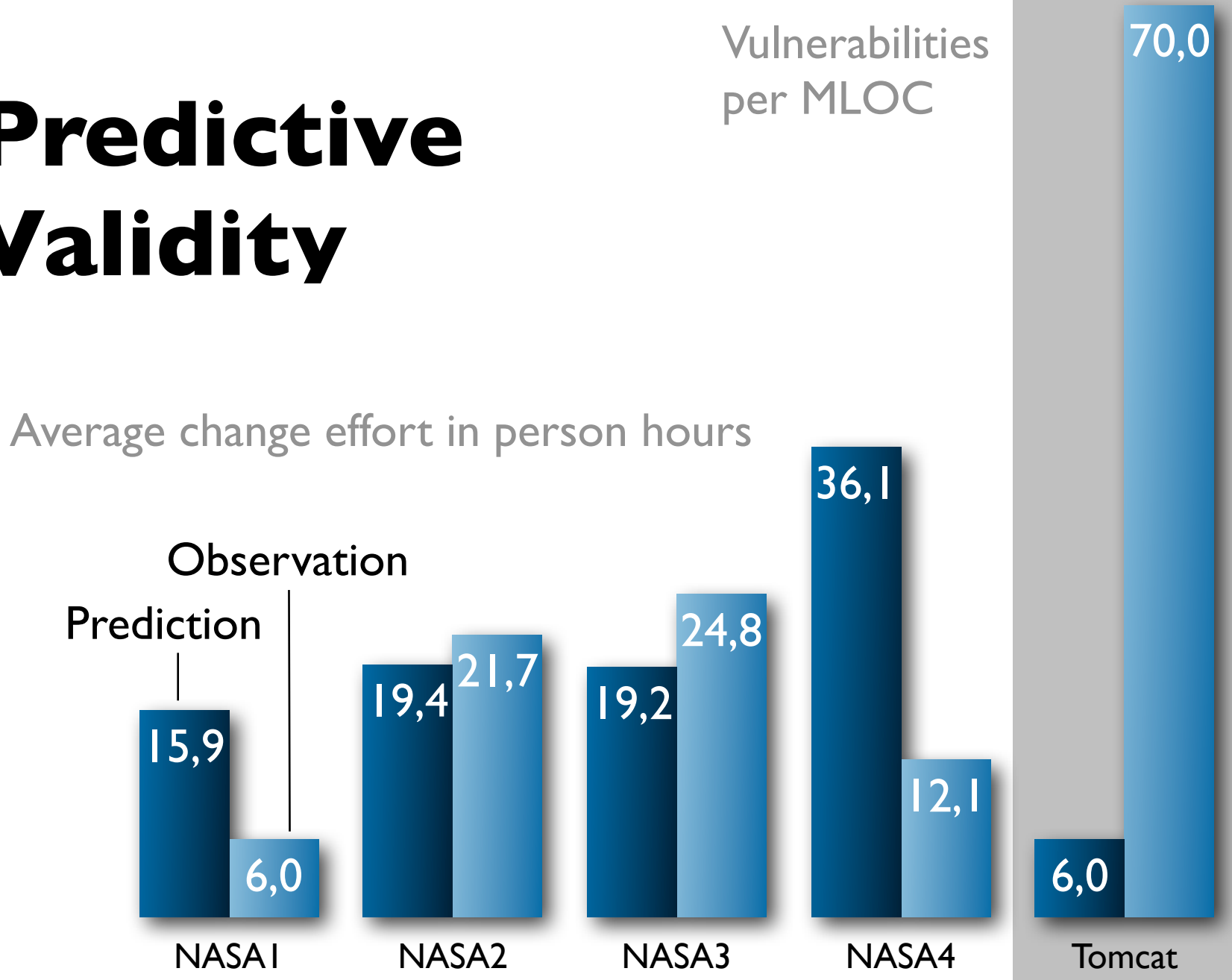
Security

Tomcat

300 KLOC
2.5 years

Wagner, Information and Software Technology, 2010

Predictive Validity



Conclusions

- Attempts to assess and predict a broader notion of quality
- Simple scoring approach
- Bayesian net approach
- Problems
  - Missing measures for quality attributes
  - Missing independent quality assessments for comparisons
  - Missing data
  - Aggregation and weighting